Imaging of infant brain using near-infrared spectroscopy: a methodological review

Jussi H. Kasurinen*[1,2], Ambika Maria[1,2,7], Riikka Korja[1,10], ### shared contribution: Anni Copeland[1,2], Venla Kumpulainen[1,2], Elmo Pulli[1,2], Eero Silver[1,2] ###, Shashank Shekhar[3], Ilkka Nissilä[4,5,9], Minna Huotilainen[1,4,5], Pauliina Hirvi[4,5,9], Linnea Karlsson[1,2,6,7], Hasse Karlsson[1,2,6,7], Jetro J. Tuulari[1,2,7,8]

Author affiliations

1 FinnBrain Birth Cohort Study, Turku Brain and Mind Center, University of Turku, Turku, Finland.

2 Department of Psychiatry, Turku University Hospital and University of Turku, Turku, Finland.

3 Department of Neurology, University of Mississippi Medical Center, Jackson, MS.

4 Department of Neuroscience and Biomedical Engineering, Aalto University, School of Science, Espoo, Finland.

5 Cognitive Brain Research Unit and Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland.

6 Department of Child Psychiatry, Turku University Hospital and University of Turku, Turku, Finland.

7 Turku Collegium for Science and Medicine (TCSM), University of Turku, Turku, Finland.

8 Department of Psychiatry, University of Oxford, Oxford, UK (Sigrid Juselius Fellowship).

9 Department of Mathematics and Systems Analysis, Aalto University, School of Science, Espoo, Finland.

10 Department of psychology, University of Turku, Turku, Finland

Corresponding author:

Jussi Kasurinen

Lemminäkäisenkatu 3, 2nd Floor, 21520, Turku, Finland

E-mail: juhekas@utu.fi

ABSTRACT

NIRS (near-infrared spectroscopy) is a relatively new neuroimaging method that utilizes the attenuation of near-infrared light within tissues and can be used to measure the changes in the concentration of oxygenated and deoxygenated hemoglobin (HbO_2 and HbR, respectively) in the tissue. It is a non-invasive method that does not involve strict movement or environment/facility restrictions and can be utilized in many age groups. Due to the challenges of data acquisition in infants and children that other neuroimaging methods (e.g. MRI) possess, NIRS is likely to be one of the key methods to study the developing brain.

This review critically complements previous reviews by summarizing the experimental set-up and probe characteristics while highlighting some of the most common problems encountered in NIRS studies on sensory processing (visual, auditory and tactile) in infants. We suggest some ways of making the best procedural choices in future studies to come. We hope that the review will help investigators compare and optimize their experimental set-ups.

Key words: NIRS; Near-infrared spectroscopy; infant; toddler; children; sensory processing; neuroimaging; systematic review

INTRODUCTION

NIRS (near-infrared spectroscopy) is a non-invasive in-vivo imaging method that detects hemodynamic changes in the brain tissue (Lloyd-Fox et al. 2010). Its use in functional neuroimaging of infants has swiftly expanded during the last twenty years. NIRS is a suitable imaging method for children because it doesn´t require strict movement restrictions and can be made in a relaxed environment, for example, while the infant is sitting on the parent´s lap. (Lloyd-Fox et al. 2010.) The temporal resolution that can be achieved by NIRS is defined by duration of the brain´s hemodynamic response, which ranges in seconds. It is better compared to functional magnetic resonance imaging (fMRI) but worse compared to electroencephalography (EEG) (Maria et al. 2018, Fig.3). The spatial resolution is largely dependent on the amount of source—detector pairs and their positioning, but NIRS is typically superior to EEG, while not as precise as fMRI. (Minagawa-Kawai et al. 2007.) Consequently, NIRS provides an excellent compromise when we want a good temporal and spatial resolution simultaneously. It is a useful method for imaging brain function in adults (Paol Pinti et al. 2018; Soltanlou et al. 2018; Ehlis et al. 2014) and in small children. (Meek 2002).

After conducting a prior review on the findings of emotional processing in infants (Maria et al. 2018), we noted that the methods were highly variable in between studies. Another key notion was that many infant NIRS studies have rather high drop-out rates (up to over 50 %) due to procedural and technical problems in the imaging session. For the most part, these problems are not only in a specific relation to NIRS imaging but are also more global relating to neuroimaging of children. The nature of the problems causing exclusion differed considering the age of the children. Older children seem to be more easily affected by these issues compared to new-born babies. Other extant reviews have addressed the available NIRS equipment (Lloyd-Fox et al. 2010; Minagawa-Kawai et al. 2007), importance of experimental designs (Bendall et al. 2016; Issard & Gervain 2018), and specifically the interpretation of positive and negative hemodynamic changes particular to infant studies (Issard & Gervain 2018).

The aim of this review is to provide an overview that enables an easier comparison between studies, equipment, protocols and analysis. Compared to our earlier review Maria et al. (2018), this review aims to provide more profound insight about other subtypes NIRS sensory processing studies, in addition to emotional processing. The sensory modalities of interest in this review are visual, auditory and tactile modalities. We aim to depict in more detail the most common problems considering these modalities that cause exclusion of study subjects. We also hope that the review will enable future investigators to make even better procedural choices when studying the sensory processing in infants, to maximize their data reliability and success rates.

METHODS

The review protocol used to review the articles and extract relevant data was similar as explained in Maria et al. (2018). We focused on sensory processing in 0 – 2 year olds in line with our prior review that specifically addressed the developmental implications of the current literature (Maria et al. 2018). The present review was done based on studies utilizing NIRS devices for assessing activation in the brain tissue. A systematic internet search of articles was made from https://www.ncbi.nlm.nih.gov/pubmed/. The search was made on September 19th 2019. We used "NIRS AND (infants OR newborns OR children) AND (brain OR neuroimaging)" as the search phrase. This search resulted in 564 items. The results were narrowed down by choosing articles containing keywords NIRS and auditory, visual, tactile, emotion, somatosensory, stimuli or pain. From the remaining articles, 27 articles that were suitable for the purpose of this review were used.

We have included articles that investigated the sensory processing of a certain modality or a combination of them. The modalities of interest in this review are visual, auditory and tactile. We were primarily interested in the early developmental aspects of the brain, and thus only the articles

that studied healthy children under the age of two years were used in this review. After choosing the suitable articles, we continued to analyse them in various aspects.

The analysis consisted of several steps. First, the studies were identified on the basis of the primary sensory modality studied and divided into three sub-groups, each consisting of articles studying either visual, auditory or tactile modality. Second, qualitative analysis about the methods used in the studies, NIRS probes, study outcomes, and the problems encountered was made for each of the sub-groups and the individual articles within the sub-groups. This enabled us to continue by assessing the differences and similarities within and between the sub-groups.

OVERVIEW OF THE REVIEWED STUDIES

Stimulus and the response

Typically, NIRS studies use block designs or event-related designs for the stimulus presentation (Lloyd-Fox et al. 2010). In a block design structure, the stimulus is repeatedly applied in blocks of a predefined temporal length. After the stimulus, there is an inter-stimulus interval (ISI), during which the subject is often presented with a neutral baseline stimulus, or no stimulus at all. In a block design, the differences between the stimulus and inter-stimulus conditions can be assessed (Cui & Bray 2012). In event-related designs, the stimuli are presented separately. The presentation is often more random as compared to a blocked design and the time between the stimuli is usually varied. Although, the ISI can also be varied in a blocked design (as in, Kotilahti et al. 2010), event related designs model the changes associated with a certain event. (Schroeter et al. 2002.)

After the stimulus is presented, the hemodynamic response can be measured. The response emerges after a lag of approximately a few seconds (Urakawa et al. 2015). Ideally, the lag before the

118 presentation of the next stimulus should be sufficiently long so that the response has time to return to
119 the baseline. In practice, the hemodynamic response´s return can take an impractically long period of
120 time, and thus, to acquire a sufficient number of repetitions of each stimulus type during a session, a
121 compromise may have to be made regarding the shortest ISI. A straightforward averaging (calculation
122 of the mean) of the response time courses over the stimulus repetitions is the most commonly used
123 method, but if there is some overlap between responses to consecutive stimuli or stimulus blocks,
124 *deconvolution* should be used to estimate the time course of the response to each stimulus condition
125 (Makni et al. 2009). The methodological designs of the studies included in this review are summarised
126 in Table 1.

127

128 ***Types of visual stimuli and responses***

129 Strong luminous light that can penetrate the eyelids has been used to assess hemodynamic changes to
130 visual stimuli in the newborn brain (as in Liao et al. 2010). More complex features of the visual
131 processing such as the processing of different shapes, colours or faces have only been studied with
132 slightly older, at least 3-5 months old, infants. Indeed, getting a neonate to concentrate to a certain
133 visual stimulus, for example a picture on a screen, may be difficult or impossible. Most studies using
134 visual stimuli have been conducted with infants over the mean age of five months (Yang et al. 2016;
135 Urakawa et al. 2015; Nakato et al. 2011; Honda et al. 2010; Wilcox et al. 2007; Liao et al. 2010;
136 Wilcox et al. 2008; Wilcox et al. 2013; Ichikawa et al. 2018), including features of visual processing
137 that have been studied in slightly older infants such as, categorical colour processing, ability to
138 differentiate between colours (Yang et al. 2016), object processing, (Wilcox et al. 2007; Wilcox et al.
139 2008; Wilcox et al. 2013) and the processing of faces and gestures (Nakato et al. 2011; Honda et al.
140 2010; Kobayashi et al. 2014; Urakawa et al. 2015; Ichikawa et al. 2018).

141

Types of auditory stimuli and responses

There are also a number of NIRS studies investigating the development of auditory processing in infants of various ages. Most of them have concentrated on how the processing of speech, language and their emotional content develops. Examples include the processing of pure tones that can be modulated on their temporal and spectral complexity (Minagawa-Kawai et al. 2011; Telkemeyer 2011), as well as different properties of speech, its processing, recognizing it´s emotional content and discriminating it from other sounds such as music (Kotilahti et al. 2010; Fava et al. 2015; Issard & Gervain 2016; Shekhar et al. 2019; Yuri Saito et al. 2007). Audio-visual stimuli can also be used to help maintain the infants´ attention during silent baseline intervals (Bortfeld et al. 2010; Bortfeld & Wruck 2012).

Types of tactile stimuli and responses

NIRS has been utilized to study different aspects of the tactile modality. Verriotis et al. (2016) studied the somatosensory responses evoked by pain and neutral touch with healthy neonate study subjects. The processing of affective touch in infants has been investigated in more depth (Kida & Shinohara 2013; Jönsson et al. 2018; Miguel et al. 2019; Pirazzoli et al. 2019). The methods used to study pain presumably differ from the methods used to study affective touch, since the stimuli studied are so different (Bennett et al. 2014) .

Table 1 summarizes the studies included in this review.

Study	Subject age (mean if reported)	Stimulus	Main objective	Main result
Visual studies				
Ichikawa et al. (2018)	3 m - 8 m, longitudinal study.	Pictures of frontal faces and right-sided profile faces. Baseline consisted of vegetables.	Investigate the longitudinal development of processing of profile and frontal **faces** in 3-8 m infants.	The processing of profile faces emerged around the age of 5 m.
Yang et al. (2016)	182.3 d	Different geometric shapes of either two shades of green or green and blue presented at 1 Hz frequency for 10 s. Baseline stimulus consisted of grey shapes. (visual)	Investigate if categorically different **colours** evoke a different hemodynamic response than categorically similar colours in the occipital and occipito-temporal areas.	Categorically different colours evoked greater hemodynamic responses in occipito-temporal areas.
Urakawa et al. (2014)	211 d	"Peek-a-boo"-game with either direct or averted gaze presented by the experimenter.	Study the effect of directed **gaze interaction** to frontal hemodynamic responses.	Hemodynamic response was greater for the directed gaze condition in the dorsomedial prefrontal cortex. Both gaze conditions caused hemodynamic responses on the right dorsolateral prefrontal cortex.
Kobayashi et al. (2014)	165.3 d and 231.8 d	10 s stimuli of flashing pictures of faces making different gestures. In the "same face" condition, the identity of the face remained the same, in the "different face" condition, the identity of the face and the gesture both changed.	Investigate how processing of **facial identity** differs between 5-6 m and 7-8 m infants.	7-8 m infants showed significantly greater hemodynamic responses in bilateral temporal areas for the different face condition. In 5-6 m, the hemodynamic responses weren´t as clearly correlated to the different stimuli.
Wilcox et al. (2012)	5 m 8 d and 11 m 21d	Shape difference, colour difference and control events shown in a puppet stage apparatus.	Investigate the differences in the processing of **colour difference and shape difference** in 3-5 m and 11-12 m infants.	Oxygenated hemoglobin (HbO_2) response was recorded in the visual cortex (occipital regions) for all events in both age groups. In older infants, a hemodynamic response was recorded also in anterior temporal areas for both color and shape difference. In younger infants, parietal cortex and anterior temporal areas showed responses for only shape difference.
Nakato et al. (2011)	200.4 d	Japanese females each posing neutral, happy, and angry facial expressions; baseline period consisted of pictures of vegetables.	Investigate the role of the temporal area (STS) in the processing of **facial expressions**.	Happy faces evoked more persisting left-lateralized hemodynamic response, while angry faces evoked a shorter left-lateralized response.
Liao et al. (2010)	2 d	Black and white board patterns shown for 1 s every 2 s in 10 s intervals. Stimulus intervals followed with 20 s baseline period of only black screen.	Investigate the hemodynamic response on the visual cortex to **simple visual stimuli** in neonates.	Most neonates showed appropriate hemodynamic responses in the visual cortex. Visual stimulus did not evoke response in the motor cortex.
Honda et al. (2009)	225.4 d	5 s visual stimulus with canonical or scrambled faces. Between different test stimulus intervals 10 s intervals of vegetable pictures as baseline.	Study the differences in the hemodynamic responses evoked by **canonical and scrambled faces**. Also compare differences in these responses between infant and adult subjects.	Total hemoglobin (Hb) concentration was significantly increased for the canonical face condition in both hemispheres. Deoxy Hb (HbR) was significantly increased in the left for the scrambled face condition. Infant responses differed from adults.

Wilcox et al. (2008)	6 m 17 d	Shape and colour difference, colour difference, shape difference and control (green ball only). Ball and box used as shapes, red and green as colours.	Investigate the hemodynamic response in inferior temporal and occipital areas for differences in **object shape and colour**.	Significant HbO2 response was recorded in occipital area for all test conditions. In inferior temporal area Significant HbO2 response was recorded for shape change and shape and colour change.
Wilcox et al. (2007)	6 m 12 d	Ball-box event, where a ball and a box were alternately presented to the infant. The ball and box differed from each other in shape and colour. Baseline consisted of no stimulus presented.	Investigate the infant hemodynamic response for processing the **object shape and colour** differences.	In the visual cortex, HbO_2 and HbT increased and HbR decreased. In the temporal region, HbO_2, HbT and HbR increased.
Auditory studies				
Shekhar et al. (2019)	55 d	11 s blocks of happy, angry, sad or neutral speech read by a Finnish actress in Finnish. Randomized 20-30 s rest period.	Investigate the neural processing of **emotional speech** in the left hemisphere.	Multiple associations found between the condition and brain areas, suggesting that the left STS is more sensitive to happy speech compared to angry, and happy speech seems to elict higher hemodynamic responses than neutral speech in the temporoparietal cortex.
Ujiie et al. (2018)	168 d	Audiovisual stimulus: wooden sound, wooden visual stimulus, metal sound and metal visual stimulus. Audio stimuli matched and mismatched with appropriate visual stimuli.	Examine the hemodynamic responses in the temporal regions evoked by **multisensory processing of different materials** in two different age groups.	The matched wood condition evoked significant hemodynamic responses in both the 4-5 m and 6-8 m groups. The matched metal condition caused significant responses in the right temporal area only in the older age group. No significant hemodynamic responses were observed for the mismatch condition nor in the left hemisphere.
Issard et al. (2016)	2.3 d	Normal speech and time compressed speech to either 60% (low compression) or 30% (high compression) from original length.	Investigate the temporal and fronto-temporal hemodynamic responses evoked by different **time compressions of speech** stimuli.	Normal speech and compression to 60% evoked a similar positive hemodynamic response, compression to 30% evoked a stronger negative response.
Fava et al. (2015)	4 m - 11 m, mean not reported	2 conditions of audiovisual stimuli. Audio stimulus was either speech (story read in infant directed intonation) or music (classical piano) depending on the condition. Visual stimulus was slowly moving shapes.	Study the differences in the hemodynamic responses on the left and right temporal areas in response to either **speech or music**.	Older infants showed larger hemodynamic responses to both of the stimulus conditions. Speech and music did not evoke significantly different responses, but there was lot of variance between subjects.
Bortfeld et al. (2012)	7 m 11 d	Audiovisual and visual stimulus conditions. The visual stimulus for both conditions consisted of pictures of slowly moving shapes, the audio stimulus was a story read in infant directed intonation. Depending on condition, both auditory and visual or only the visual stimulus were used simultaneously.	Examine the hemodynamic response for **visual and audioviusual stimuli** in the left temporal and occipital area.	Both conditions evoked a significant hemodynamic response in the occipital area while only the audiovisual stimulus evoked a response in the temporal area.

Telkemeyer et al. (2011)	Group 1: 185 d, group 2: 94 d	Temporally modulated sounds (slow and fast) were presented.	Study how the processing of **temporally fast and slow sounds** develop in different age groups.	Older group showed larger hemodynamic responses for the temporally fast sounds in the left temporal region for the temporally slow sounds in the right temporal region. Younger group showed larger hemodynamic responses on the left side for both conditions.
Minagawa-Kawai et al. (2011)	2.4 d	Three pure tone patterns that varied on spectral and temporal properties. The control pattern was spectrally and temporally simple while the other stimuli were either spectrally or temporally complex.	Test the differences in the hemodynamic responses for the different **spectral and temporal properties of sounds**.	Temporally complex sounds evoked significantly greater responses on bilateral temporal areas compared to control or spectrally complex stimuli. No significant lateralization was observed.
Taga et al. (2011)	111 d	2 conditions of 5 s auditory stimulus of 25 spectrally different random puretones of 100 ms in duration. Interstimulus intervals were different in the 2 conditions, either 10 s or 5 s.	Investigate the properties of the cortical hemodynamic response and the timing of the cortical activation flow in response to **simple auditory stimuli**.	Wide cortical activation was detected when the interstimulus interval was longer 10 s. The patterns were dissociated in the 5 s interstimulus interval condition.
Kotilahti et al. (2010)	1.8 d	5 s samples of speech (story read using infant directed intonation) and music (Mozart piano concerto). Varying interstimulus interval (mean 15 s). Stimuli presented up to 64 times. (auditory)	Investigate the processing of **speech and music** on newborns.	Both hemispheres could be activated by both stimuli although there were individual variations. No significant lateralization was observed for either of the conditions. Speech evoked a more uniform response on the left hemisphere.
Bortfeld et al. (2009)	6 m - 9 m, mean not reported	Audiovisual and visual stimulus conditions. Visual stimulus for both conditions were pictures of slowly moving shapes, audio stimulus was a story read in infant directed intonation. Depending on condition, both auditory and visual or only the visual stimulus were used simultaneously.	Examine the laterality of the hemodynamic response in the temporal areas for **audiovisual and visual stimuli**.	Audiovisual stimulus evoked a significantly greater response in the left temporal region.
Saito et al. (2006)	4.7 d	Speech stimulus where a story was read in a female voice. The stimulus was digitally altered to be spectrally monotonous or variable. Stimuli were 30 s long.	Investigate the differences in the processing of **variable and monotonous speech** in neonates.	Variable speech produced a greater oxygenated hemoglobin (HbO_2) response in the frontal areas than monotonous speech.
Saito et al. (2006)	4.4 d	Story read by the infants own mother in infant directed speech or adult directed speech. White noise used as the baseline stimulus.	Examine the relation of hemodynamic activity in the frontal areas of the neonatal brain and **maternal infant directed speech**.	Maternal infant directed speech caused greater hemodynamic responses in the frontal areas compared to maternal adult directed speech.
Tactile studies				

Miguel et al. (2019)	228.7 d	Infant´s right dorsal forearm was stroked slowly (8 cm/s) using a wide watercolor brush for producing the affective stimulus and square shaped piece of wood for the discriminative stimulus. The discriminative stimulus did not include stroking movement.	Investigate the hemodynamic responses for **affective touch** in the somatosensory areas and temporal regions.	Both stimuli evoked similar significant responses in the corresponding somatosensory cortex. No activation was registered in the temporal region.
Pirazzoli et al. (2019)	160.2 d	Gentle stroking (3 - 10 cm/s) to infants forearm, applied with either the experimenters hand or a spoon.	Investigate the hemodynamic responses for **affective touch** in the posterior superioir temporal gyri and inferior frontal gyri.	The responses in the studied brain areas did not differ between the two stimulus conditions.
Jönsson et al. (2018)	56 d	Infant´s arm was stroked gently with a soft goat hair brush at two velocities 3 and 20 cm/s.	Investigate the hemodynamic responses for **affective touch** in the temporal areas.	Carressing touch evoked significantly greater hemodynamic responses in the insula and middle temporal gyrus compared to fast touch.
Verriotis et al. (2016)	2 d	Noxious somatosensory stimulus (heel lance), innocuous tactile stimulus (touch) and control stimulus (heel lance, no sample taken).	Examine the relationship between newborn EEG and NIRS responses for **noxious somatosensory stimulus**.	Noxious stimulus evoked significantly greater hemodynamic response on the contralateral somatosensory cortex compared to the innocious stimuli. The NIRS and EEG responses were consistent in 64% of cases.
Kida et al. (2013)	6.6 m	Affective touch (infants palm stroke with a velvet wrapped cotton packed piece of wood) andneutral touch (column shaped piece of wood used for stroking).	Investigate the hemodynamic response for **affective touch** in the anterior prefrontal cortex in 3, 6, and 10 month old infants.	The 10 month old group showed a significantly greater hemodynamic response for gentle touch compared to neutral touch in the bilateral anterior prefrontal cortex, while the younger age groups didn´t show significantly different hemodynamic responses for the different touch conditions.

Table 1. The main objective, stimuli and results used in the reviewed studies. Abbreviations: m: month, d: day, s: second.

DATA ACQUSITION

The probe

The NIRS probe consists of emitters and detectors that are held against the scalp of the subject over the area of interest. It is held in place using a head band or a wearable cap-type system. The probe and head gear designs have been developed to be applied easily and comfortably to the infant´s head

because using uncomfortable probes that are hard to apply to the infant can result in experiment failure. The comfortability and accessibility of the probes are important matters to be considered when designing them.

Infants are not easily co-operative to the experiment if they find the probe uncomfortable or if the setting up of probe onto the infant´s head takes too long. Thus, the practical design of the probe used influences the success rate, especially in infant studies. Sophisticated, easily applicable probe designs provide better study outcomes in terms of the number of successful experiments. (Lloyd-Fox et al. 2010.) To maximize the comfortability of the probe, silicon probe holders and pads have been developed to keep the probe on the area of interest (as in Kobayashi et al. 2014; Nakato et al. 2011; Liao et al. 2010; Kobayashi et al. 2011). Ongoing development of the probe and the systems that are used to hold it in place will hopefully provide us with even more easily applicable probes in the future.

Also, commercial NIRS imaging devices have been developed. Hamamatsu Photonics has come up with a commercial NIRS system called the NIRO (http://www.hamamatsu.com/eu/en/index.html?gclid=COrigKb9uNQCFR1lGQodZhMB2w) (Y Saito et al. 2007; Ranger et al. 2011; Yuri Saito et al. 2007; Verriotis et al. 2016; Kida & Shinohara 2013). ISS has developed another commercial NIRS instrument called the Imagent (http://www.iss.com/biomedical/instruments/imagent.html). Apart of the above mentioned, one of the most widespread systems is the ETG optical topography system by Hitachi Medical Corporation. (http://www.hitachi.com/businesses/healthcare/products-support/opt/index.html). The ETG system has been utilized in many studies to investigate infants´ modal processing (Yang et al. 2016; Honda et al. 2010; Kobayashi et al. 2014; Ichikawa et al. 2014; Taga et al. 2011; Nakato et al. 2011; Kobayashi et al. 2011). Using the same commercial imaging system can be postulated to be a factor

194 positively affecting reliability of inter-study comparisons in some situations. In addition, there are a
195 variety of other commercial instrumentations that may also incorporate other imaging methods with
196 NIRS. For example, CerOx, c-Flow by Ornim medical Israel (http://www.ornim.com/) combines
197 ultrasound imaging with NIRS and provides a way of assessing hemoglobin concentrations and blood
198 flow simultaneously. Unfortunately, the device doesn't come with an infant sensor. (Sood et al. 2015.)
199 For more details: please refer to the manufacturers for the latest available equipment, also reviewed
200 more extensively in Scholkmann et al. (2014).

201

202 ***Placement of the probe***

203 The NIRS imaging technique is based on measuring the (changes in the) attenuation of near-infrared
204 light in the tissue between the light source or emitter and the detector, which are placed a few
205 centimeters apart on the scalp. One detector-emitter pair forms one channel and in order to acquire
206 2D optical topography or 3D optical tomography images, NIRS studies utilize multiple channels to
207 localize the response and form an image of the brain's hemodynamic activity during the response.
208 Although, also using only a single channel can be utilized to assess local hemoglobin concentration
209 changes, multichannel measurements provide us with an more accurate conception of the local
210 hemodynamic responses properties. (Lloyd-Fox et al. 2010)

211

212 The 10-20 international EEG system is a commonly used method to describe the location of the NIRS
213 optodes on the scalp. The positions of the international EEG system are based on external marks on
214 the scalp (Okamoto et al. 2004). The individual differences in the shape of the head can affect the
215 accuracy of the hemoglobin concentration measurements in terms of both, localization and
216 quantitative accuracy (Sood et al. 2015). Compared to EEG studies, most NIRS studies investigate
217 only a portion of the brain´s cortex and thus only a fraction of the possible electrode placements

218 described by the 10-20 international EEG system are used simultaneously. Also, other likely more
219 precise ways of navigating the optode placement such as using infant MRI-templates or individual
220 MRI-pictures coupled with photogrammetry have been implemented. (Heiskala et al. 2009; Ferradal
221 et al. 2014; Shekhar et al. 2019).

222

223 A NIRS device can have multiple detector-emitter pairs that simultaneously measure from different
224 areas. Thus, a larger portion of the brain´s cortex can be assessed simultaneously. Probes covering
225 extensive areas of the infant´s brain can provide the researchers with a broader picture of the brain
226 activation considering a certain stimulus (as in Taga et al. 2011), while some studies utilize only one
227 or two channels to assess the hemodynamic changes in a more concise area (as in Kida & Shinohara
228 2013). Generally, the layout of the detectors and the emitters on the probe influences the areas that
229 can be monitored, the resolution and aspects of the reliability and reproducibility of the data (Heiskala
230 et al. 2009).

231

232 The sensitivity of the measurement to physiological changes occurring at different depths can be
233 adjusted to some extent by changing the distance between the detector and the emitter. A longer
234 emitter–detector distance increases the relative contribution of light that has propagates through
235 deeper tissue to the signal corresponding to that channel, as shown in Fig. 2. The visualized sensitivity
236 profiles (or absorption Jacobians) were computed with the Monte Carlo eXtreme software (Fang et
237 al. 2009, Yao et al. 2018) with emitter–detector distances of 2.3 mm (S1–D1, short channel) and 25
238 mm (S2–D2, long channel) (Brigadoi et al. 2015). The MR-image is of a 1.5-month-old infant from
239 the FinnBrain study in Hashempour et al. (2019) with tissue-wise optical parameters as in Jönsson et
240 al. (2018). In addition, a combination of channels of different lengths can be used to assess
241 hemodynamic changes from deeper and more superficial structures simultaneously, using the same
242 probe (as in, Minagawa-Kawai 2011). However, it has been argued that due to optical properties of

brain tissue, namely the white matter, the quality of measurement weakens when the source-detector distance is increased. In addition to having smaller heads and thinner scalp and skull, neonates have also lower density in white matter compared to adults, which facilitates the measurement and imaging of processes taking place in deeper parts of the cortex. (Fukui et al. 2003.) There is no clear consensus one most appropriate detector-emitter distance, the distances vary between studies and rather the best way is to use a range of distances to collect as much information as possible (Jönsson et al. 2018; Shekhar et al. 2019). However, an approximately 2-cm source-detector distance has been commonly used in infant studies to image the infant cortex. (see Table 2.)

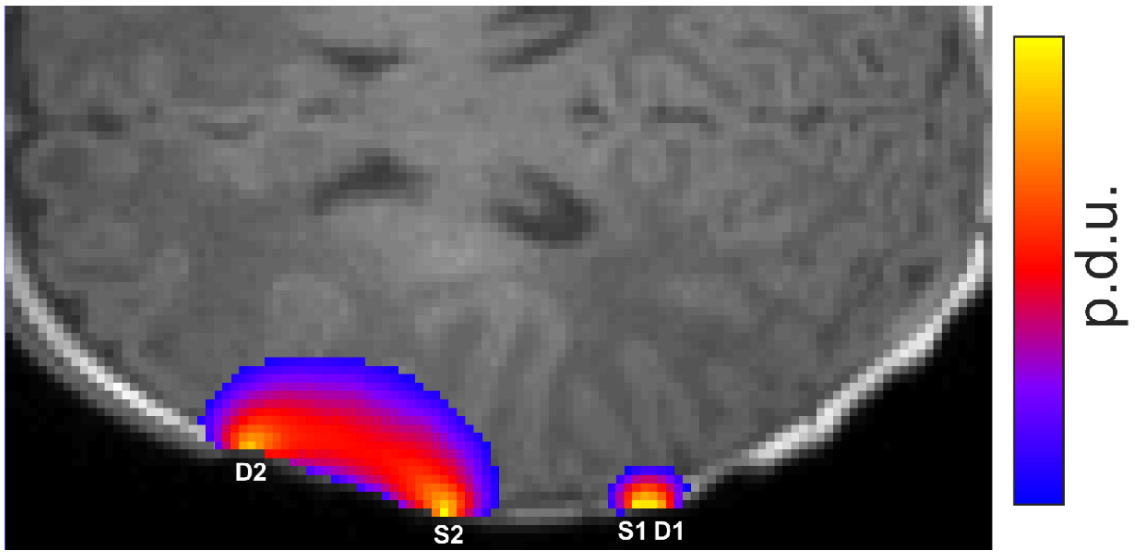
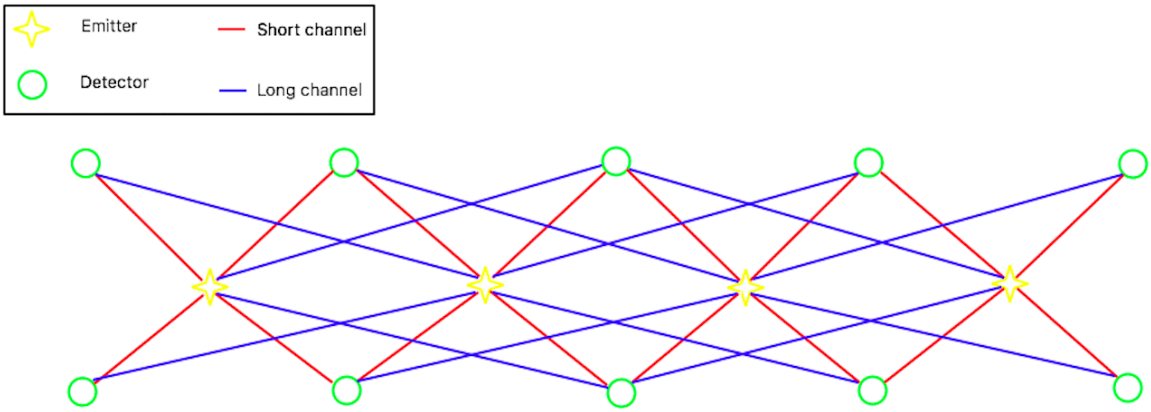


Figure 1. The short channel between source S1 and detector D1 is mainly limited to extracerebral regions. The longer emitter–detector distance between source S2 and detector D2 allows deeper structures to be imaged. However, there are limitations due to the diffuse light reflection between white and grey matter. The cerebrospinal fluid also tends to concentrate sensitivity in its vicinity (Heiskala et al. 2009).

NIRS offers rather good temporal resolution that is superior to fMRI and PET but not as good as EEG or magnetoencephalography (MEG). While the NIRS signal can be measured at a very fast rate, 10 Hz is an imaging frequency often used. The underlying physiological changes limit the temporal resolution of genuine physiological features that can be extracted from the signal that are expected to have a relationship with the underlying neuronal activity. Also, slightly lower frequencies such as 6 Hz have been utilized (as in Kida & Shinohara 2013). Similar converse between NIRS and other neuroimaging methods is true of the spatial resolution; fMRI offers higher spatial resolution compared to NIRS and there is no ill-posed inverse problem in fMRI. On the other hand, the spatial resolution of NIRS is superior to that of EEG. (Minagawa-Kawai et al. 2007). Spatial resolution can be characterized as lateral and depth resolution.

Lateral and depth resolution in NIRS imaging are determined by the differences in the optode arrangements. Different separations between the emitters and detectors, and the fibre arrangements, where one detector detects light from multiple emitters (multi-distance measurements) offers improvements in the depth resolution (Liao et al. 2010). Using different channel arrangements can enhance the spatial resolution and quantitative accuracy achieved. High-density fibre arrangements have been shown to have an effective resolutions of such as 1 cm, which is enough to identify gyri on the brain´s cortex (White & Culver 2010).



277

278 *Figure 2 A schematic presentation of multi-distance measurement, where four emitters and ten*
279 *detectors form a total of 16 short channels and 12 long channels. Similar principle is used to construct*
280 *the high-density arrangements mentioned above.*

281

Study	Probe used	Number of channels	Number of optodes	Source-detector separation	Placement of probe
Ichikawa et al. (2018)	ETG-4000; Hitachi Medical	24 channels total, 2x 3x3 arrays. 12 channels left, 12 right.	10 sources, 8 detectors total.	2 cm	Probes placed over bilateral temporal areas T5 and T6 according to the international 10-20 system.
Yang et al. (2016)	ETG-4000; Hitachi Medical	24 channels with 2x 3x3 arrays of sources. In addition,occipital measurement with 12 channels used 1x 3x3 array	10 sources, 8 detectors (5 sources and 4 detectors on 1 probe)	2 cm	Probes placed slightly below T5 and T6 using the international 10-20 system for the occipitotemporal areas. For additional occipital measurement, the probe was placed slightly above Oz.
Urakawa et al. (2014)	OMM 3000, Shimadzu Corporation	17 channels	6 emitters, 6 detectors.	2 cm	Most ventral probes were placed at Fp1 and Fp2.
Kobayashi et al. (2014)	ETG-4000; Hitachi Medical	24 channels, 2x 3x3 arrays. 12 channels left, 12 right.	10 sources, 8 detectors total.	2 cm	Temporal region, placed at T5 and T6 according to the international 10-20 system.
Wilcox et al. (2012)	NA	9 channels, 1x L-shaped array on left side.	4 sources, 8 detectors.	2 cm	Left side parietal, temporal and ooccipital regions. Emitters placed at T3, T5, O1 and P3.

Nakato et al. (2011)	ETG-4000; Hitachi Medical	24 channels, 2x 3x3 arrays. 12 channels left, 12 right.	10 sources, 8 detectors total.	2 cm	Central position of probes were at T5 and T6 according to the international 10-20 system.
Liao et al. (2010)	Self assembled multichannel continuos wave high density NIRS device.	Number of channels not reported, long and short channel system in a 4x10 configuration.	4 sources, 10 detectors	short channel 1 cm, long channel 2.3 cm	Inion used as a landmark for the visual cortex, probe placed over the inion.
Honda et al. (2009)	ETG-100; Hitachi Medical	24 channels, 2x source detector arrays on each side consisting of 5 emitters and 4 detectors.	10 sources, 8 detectors total	2 cm	Occipitotemporal region, placed over T5 and T6 according to international 10-20 system.
Wilcox et al. (2008)	NA	4 channels, 2x 1 emitter, 2 detector systems. 1 used on occipital area, 1 used in left inferior temporal area.	2 sources 4 detectors total.	2 cm	Left side inferior temporal area and occipital area. Emitters placed between O1 and O2 on occipitum and T3 on inferior temporal area.
Wilcox et al. (2007)	NA	2x emitter detector triads. 1 used on left temporal area, 1 on occipital.	2 sources 4 detectors total.	2 cm	Temporal emitter set at T3, occipital cortex, directly above the inion.
Shekhar et al. (2019)	Custom diffuse optical tomography (DOT) system built at Aalto University.	High-density system with multiple channels of different length. Total number of channels 160.	15 emitters, 15 detectors	Multiple different lengths. 7 mm to 45 mm used for image reconstruction.	Probe was positioned on the left side of the infant's head. The size of the infant´s head was measured and sterephotography with multiple sticker markers was used for constructing a 3D model of the infant´s head with anatomical land marks that were subsequently used for determining the position of the probe.
Ujiie et al. (2018)	ETG-4000; Hitachi Medical	24 channels total, 2x 3x3 arrays. 12 channels left, 12 right.	10 sources, 8 detectors (5 sources and 4 detectors on 1 probe)	2 cm	Probes placed at T3 and T4 according to the international 10-20 system.
Issard et al. (2016)	NIRScout 816; NIRx Medizintechnik GmbH	Total 24 channels, 12 on each side.	Total 8 sources, 10 detectors. 4 sources, 5 detectors on each side.	3 cm	Placed on bilateral fronto-temporal and temporal areas using newborn average MRI template to verify the localization.

Fava et al. (2015)	NA	4 channels, 2x right temporal, 2x left temporal.	Altogether 2x emitters, 4x detectors total. 1 detector 2 emitters placed over both regions of interest.	2 cm	Probes placed at T3 and T4 according to the international 10-20 system.
Bortfeld et al. (2012)	NA	4 channels, 2x occipital, 2x left temporal.	Altogether 2x emitters, 4x detectors total. 1 detector 2 emitters placed over both regions of interest.	2 cm	Occipital set placed 1-2 cm above the inion, temporal set placed slightly above and anterior to T3 according to the international 10-20 system.
Telkemeyer et al. (2011)	Omniat Tissue Oxymeter; ISS	12 channels total, 6 on each side	8x emitters, 4x detectors	2.5 cm	10-20 international system used to place the probe to cover areas in the fronto-temporo-parietal area of both hemispheres.
Minagawa-Kawai et al. (2011)	UCL-NTS, NIRS device developed in the Department of Medical Physics and Bioengineering of University College London	28 channels total. 10 superficial and 4 deep channels on each side.	8 sources and 8 detectors total, 4 on each side.	2.5 cm and 5.6 cm	Bottom of the pad aligned with the T3-T5 line according to the international 10-20 system.
Taga et al. (2011)	ETG-7000; Hitachi Medical	94 channels	10 emitter detector triple piece sets on each side.	2 cm	Arranged over occipital, temporal and frontal cortices using Fpz, T3 and T4 according to onternational 10-20 system as reference points.
Kotilahti et al. (2010)	NIRS device developed at Helsinki University of Technology.	16 channels, 8 channels on each side	2x sources, 16x detectors total.	1.5 cm and 2.5 cm	Bilateral temporal areas.
Bortfeld et al. (2009)	NA	4 channels, 2x right temporal, 2x left temporal.	2 sources and 4 detectors total.	2 cm	T3 and T4 according to international 10-20 system were used as reference points which the probe was placed on.
Saito et al. (2006)	NIRO-200; Hamamatsu Photonics	2 channels	2 sources and 2 detectors, 1 on each side	3 cm	Left and right frontal areas Fp1 and Fp2 according to internetional 10-20 system used as markers for placement.
Saito et al. (2006)	NIRO-200; Hamamatsu Photonics	2 channels	2 sources and 2 detectors, 1 on each side	did not report	Left and right frontal areas; Fp1 and Fp2 of the international 10-20 system used as markers for placement.

Miguel et al. (2019)	UCL – fNIRS to-pography system	18 channels	12 emitters, 6 detectors.	short channels 20 – 25 mm, long channels 45 mm	Easy Cap; reference 10–5 system used in the NIRS-cap. The center of the cap was placed at Cz and channel 11 at TP8.
Pirazzoli et al. (2019)	Custom built CBCD - NIRS headgear	26 channels	NA	20 mm	Inion, nasion, and ears were used as anatomical landmarks for assessing the placement of the probe.
Jönsson et al. (2018)	Custom diffuse optical tomography (DOT) system built at Aalto University.	High-density system with multiple channels of different lenght. Total number of channels 107.	NA	Multiple different lenghts ranging from 7 mm to 45 mm.	Probe was positioned on the left side of infants head, right above the ear. Stereo photogrammetry was used to register the position of the optodes with the anatomical model based on anatomical landmarks.
Verriotis et al. (2016)	NIRO-200NX; Hamamatsu Photonics	1 channel	single emitter detector pair	4 cm	Mid-point of the emitter detector pair was placed over C1 or C2 (according to international 10/20 elctrode placement system) depending on which side was contralateral for the corresponding stimulus.
Kida et al. (2013)	NIRO-200; Hamamatsu Photonics	2 channels	2 sets of 1 detector 1 emitter pairs.	NA	Fp1 and Fp2 according to international 10-20 system used as reference points, where the probe was placed.

Table 2. Properties of the probes used in the studies reviewed.

DATA ANALYSIS

The data analysis is carried out quite similarly in the NIRS studies included this review. NIRS probe captures the attenuation signal using continuous wave (CW), time-domain (TD) or frequency-domain (FD) system (Minagawa-Kawai et al. 2007). CW systems are most usually used in infant NIRS studies (Lloyd-Fox et al. 2010).

The analysis consists of a sequence of operations that help differentiate between significant hemodynamic responses related to neuronal activity, physiological variations in the hemoglobin concentrations and artefacts in the measured signal. The changes in the attenuation are transformed

into changes in the hemoglobin concentrations using the modified Beer-Lambert law. The data is then high- and low-pass filtered to remove high frequency oscillations caused, for example, by the pulsation of the heart and to remove linear or non-linear drift from the data. Both high- and lowpass filtering, if inappropriate frequencies are used, can also distort the time course of the hemodynamic response. Thus, the cutoff frequencies must be chosen carefully. Simulations can be used to assess the distortion and choose appropriate frequencies so that high- and lowpass filtering doesn´t lead to misinterpretation of the results. Large peaks caused by head movement or other artefact-causing factors are removed, commonly by using a cut-off threshold. (Lloyd-Fox et al. 2010.) In addition to the above-mentioned signal processing steps, other techniques, like visual monitoring and video monitoring or other more objective methods such as inclinometers (as in Kotilahti et al. 2010), are often used to identify potential epochs where artefacts may have occured.

An inclinometer can help to identify movement artefacts (Kotilahti et al. 2010). Likewise, video monitoring can be utilized to assist in identifying and removing movement artefacts (Jönsson et al. 2018) and to assess infants´ compliance during the presentation of the stimulus. This is a particularly advisable method in studies that use visual stimuli, because it can additionally help to identify and delete epochs where the infant´s attention is not directed to the stimulus. (Honda et al. 2010.) After the hemodynamic responses have been identified and isolated from the attenuation data, statistical analyses can be conducted to determine the significance of the responses.

Conventional statistical methods like analysis of variance (ANOVA) and Student's t-test can be utilized to determine statistical differences between the responses (Lloyd-Fox et al. 2010). However, before the statistical analysis can be applied, the data is often resampled to a desired frequency (e.g.

316 1 Hz) in the general case by interpolation, low-pass filtering and decimation (as in, Kotilahti et al.
317 2010; Hull 2015).

318

319 The filtered time course can be modelled using one or several explanatory variables with the general
320 linear model (GLM). After that, different approaches of assessing statistical significance can be used.
321 The area-under-the-curve (AUC) during the response and the baseline can be calculated and
322 compared to each other (as in Kotilahti et al. 2010). In addition to the AUC, the means of the
323 hemoglobin concentrations during these two time windows can be compared. Also, the peak values
324 of the responses are sometimes compared (as in Verriotis et al. 2016), even though this can be seen
325 as bad practice because searching for the peak of a noisy signal time course can lead to an
326 overestimation of the absolute value of the response.

327

328 *T*HE SUCCESS RATES OF THE REVIEWED STUDIES

329 The success rates of infant and neonate NIRS studies are quite commonly affected by subject
330 compliance during the measurements. Naturally, the period when the probe is fitted to the infant´s
331 head is critical for the whole experiment. If infants do not accept the probe, no data can be collected.
332 In addition, technical matters like loose attachment of optodes and hair obstruction can cause a failure
333 to obtain good quality data. These problems can often be avoided by a careful fitting of the probe and
334 a good preparation for the imaging session.

335

336 However, lack of concentration to the stimulus and signs of unease such as crying can cause large
337 movement artefacts and difficulties in obtaining good quality data. They can also be interpreted as
338 the infant's way of disapproving to take part in the study, and thus often lead to the discontinuation

of the imaging session. The headgear design affects its comfortability and can have a large effect on getting the infant subject to co-operate in the study.

Probes that have a large number of emitters and detectors are often heavier, more difficult to set up and may be less comfortable to the infant. This can lead to a larger percentage of subjects not agreeing to wear the probe. For example, Taga et al. (2011) excluded 53% of their infant subjects from the analysis on basis of not sleeping or sleeping too restlessly in their study using a relatively big 94–channel probe. The mean age of the subjects in the study was 111 days. In comparison, Hull et al. (2015) excluded only 9% of their 4 –11-month-old infant participants by using a much more compact probe of only 4 channels. Both studies were conducted using auditory stimuli. However, a four-channel recording cannot answer the same questions that a 94–channel recording can. Using larger probes with more channels can naturally provide us with larger imaging field of view (FOV), better spatial resolution and more robust results (e.g., Heiskala et al. 2009). Thus, such frank comparisons between the exclusion rates and probe size must be interpreted very carefully, and of course, there are also other matters, in addition to the probe design, that may affect the number of subjects excluded from the analysis (e.g. subject age, the complexity of the study protocol, involvement of the parents to the experiment and the type of the stimulus used).

The age of the subjects influences strongly the study compliance and exclusion rates. In general, studies with neonates that are only a few days old result in smaller exclusion rates compared to studies conducted with slightly older infants (see Tables 3, 4 and 5). Neonates naturally spend most of their time sleeping and are not yet really bothered by the environmental factors considering imaging process. Most of neonatal studies are conducted so that the subject is in natural sleep, which can even result to full success rates, so that none of the imaging sessions are not rejected from analysis (as in

Kotilahti et al. 2010; Yuri Saito et al. 2007; Y Saito et al. 2007). If there are exclusions, they are usually caused by large movement artefacts, or technical difficulties such as loose probe attachments that cause data losses (as in Liao et al. 2010; Dupoux 2011; Verrriotis et al. 2016). Overall, the percentage of subjects excluded in neonatal studies tends to be lower than in older age groups of children.

The high rates of subjects excluded from analysis in older age groups is probably mainly due to a few issues considering the age group. Older infants are not as easily put to natural sleep during day time as neonates are. Also, the experimental protocols and stimuli are often more complex, which can lead to procedural problems such as crying or inability to concentrate to the stimulus (Kobayashi et al. 2014). For example, Yang et al. (2016) used a relatively complex experimental protocol (see Table 1) to study infants´ (mean age of 182.3 d) categorical colour perception. They excluded 43% of the total participants on the basis of not looking at the stimulus or other biases. As stated, it also seems that older infants tend to sleep more restlessly than neonates. Taga et al. (2011) studied 111 day old infants during natural daytime sleep and 53% of the subjects were excluded mostly due too restless sleep or not sleeping. Also, other studies of older infant participants show a trend of higher exclusion rates compared to neonatal studies.

COMMON PROBLEMS IN SPECIFIC SENSORY MODALITIES

Infant NIRS studies conducted to investigate the processing of different modalities show differences in the study structures considering a certain sensory modality. Thus, the problems encountered during the studies can be characterized by the modality studied (visual, auditory and tactile).

Visual

Studies with visual stimuli tend have the most complicated experiment protocols. In study topics like object processing (Wilcox et al. 2013; Wilcox et al. 2007; Wilcox et al. 2008) or the processing of different facial gestures (Kobayashi et al. 2014; Nakato et al. 2011), many different visual stimuli are alternately shown to the subject, resulting in complicated protocols that also demand the subject´s performance during the experiment to be monitored.

Video monitoring the subject´s gaze is a widely used method of assessing fixation to the stimulus. It can be used during the study to identify moments when the infant is ready to concentrate to the stimulus and present the stimulus only at those moments (as in Honda et al. 2010). Nevertheless, the most common way is to use video monitoring after the experiment to identify the moments when the infant´s gaze was directed to the stimulus and only use them in the analysis (as in Yang et al. 2016). In order to get even more precise information about the subject´s looking behaviour more advanced methods such as eye tracking have also been used (as in Urakawa et al. 2015). Eye tracking was also utilized to investigate which parts of the stimulus picture presented caught the infants´ attention.

Visual stimulus is often presented on a computer screen. This causes changes in the experiment room´s luminance, which can interfere with the detectors of older NIRS device´s and cause artefacts to the data. The artefacts can be controlled by adjusting the luminance properties of the stimulus so that it doesn´t interfere with the NIRS imaging or by using more modern NIRS devices. (Yang et al. 2016; supporting infromation Control et al. n.d.). It is advisable to use modern devices that use modulated source light to separate the changes in the signal from the changes in the ambient light, when ambient light cannot be dimmed to be undetectable.

409

410 Exclusion rates in visual studies tend to be higher compared to other modalities (see Table 3). For
411 example, in Nakato et al. (2011) and Yang et al. (2016), exclusion rates in visual NIRS studies can
412 range up to 43%. Moreover exclusion rates of over 30% are not rare at all (Urakawa et al. 2015;
413 Wilcox et al. 2008; Wilcox et al. 2007). Visual attention is not as developed compared to senses of
414 touching and hearing during the infancy, and thus the high exclusion rates are typically caused by
415 challenges in getting the participant to concentrate on the stimulus.

416

Study	Subject age (mean)	Method of monitoring behaviour	Exclusion criteria	Exclusion rate
Liao et al. (2010)	2 d	NA	Exclusion of one infant because of large movement artefacts. Strong pacifier sucking was observed.	1 / 11 = 9,1%
Yang et al. (2016)	182.3 d	Video tape	Subject was excluded from analysis if the looking time for the test stimuli was less than 7 s or if they became fussy.	18 / 42 = 42.9%
Urakawa et. Al. (2014)	211 d	Video tape, eye tracking	Failure to concentrate gaze to the stimulus in more than 6/12 trials.	7 / 18 = 38.9%
Nakato et al. (2011)	200.42 d	Video tape	4 subjects because of not being able to concentrate to stimulus, 2 crying, 2 big movement artefacts, 1 computer error.	9 / 21 = 42.9%
Honda et al. (2009)	225.4 d	Video tape	Less than 3 successful trials led to exclusion. Subjects excluded because of crying, not looking at the stimuli or movement artefacts.	3 / 16 = 18.8%
Kobayashi et al. (2014)	165.3 d and 231.8 d	Video Tape	Rejection due less than three successful trials of both conditions. Exclusions on base of crying or motion artefact.	7 / 31 = 22.6%
Wilcox et al. (2012)	5 m 8 d and 11 m 21 d	2 individual observers, that watched the infant trough peep holes during the study	Failure to look at 2 or more trials at least 10 s or procedural problems, difficulty obtaining optical signal, movement and crying.	37 / 148 = 25%

Wilcox et al. (2008)	6 m 17 d	2 individual observers, that watched the infant trough peep holes during the study	Large motion artefacts in the signals (N = 6), obstruction by hair (N = 2), procedural problems (N = 4), Data available from only 1 brain region (N=6)	18 / 53 = 34.0%
Wilcox et al. (2007)	6m 12d	2 individual observers, that watched the infant trough peep holes during the study	Large motion artefacts in the signals (N = 2) or failure to obtain adequate signals because of obstruction by hair (N = 1)	3 / 10 = 30%
Ichikawa et al. (2018)	longitudinal 3 m – 8 m.	Video monitoring	Failure to look at the stimulus, if infant became fussy, movement artefacts.	NA
				Mean exclusion percentage: 29.36%

Table 3. Exclusions in visual studies. Abbreviations: m: month, d: day

Auditory

Studies investigating the development of auditory processing often have rather simple stimulus-response structures, where the auditory stimulus is usually presented to the infant subject, who passively listens to it. Assessing how directed the infant´s attention is to the stimulus can´t be done as precisely as in visual studies, but the infant´s state (e.g., calm state with eyes closed/awake) can be monitored and reported. Studies have been conducted with the subject is awake or asleep (or mixed).

Compared to visual studies discussed above, due to simpler experiment protocols, the exclusion rates tend to be markedly lower in studies using auditory stimuli. Studies seldom report exclusion rates of over 25% (see Table 4).

Study	Subject age (mean)	Method of monitoring behaviour	Exclusion criteria	Exclusion rate

Shekhar et al. (2019)	55 d	Video recording and artefacts identified on basis of this video and hemodynamic data.	Head movement, crying or limb movement that caused insufficient number (<5) of artifact-free trials.	25/46 = 54.3%
Ujiie et al. (2018)	168 d	Digital recording behaviour throughout the experiment.	Fussiness, motion artefacts that caused insufficient number (4) of trials.	7/39 = 17.9%
Issard et al. (2016)	2.34 d	Visual observing during experiment.	Crying, technical problems and movement artefacts hair obstruction.	38/59 = 64.4%
Fava et al. (2015)	4 m – 11 m, mean not reported	Video recorded, behaviour assessed on base of these recordings.	No useable blocks of data were obtained or no optical data was collected for both conditions.	4 / 45 = 8.9%
Bortfeld et al. (2012)	7 m 11d	Video recorded, behaviour assessed on base of these recordings.	Movement artefacts, hair obstruction or not able to obtain a useful block of data.	5 / 40 = 12.5%
Taga et al. (2011)	111 d	Video recorded, artefacts identified on basis of video and hemodynamic data.	Not staying asleep throughout the study or sleeping too restlessly.	20 / 38 = 52.6%
Telkemeyer et al. (2011)	Group 1: 185 d, group 2: 94 d	Visual observing during experiment.	Movement artefacts, infant or parent discomfort.	13 / 84 = 15.5%
Minagawa-Kawai et al. (2011)	2.41 d	Visual observing during experiment.	interruption due to infant discomfort, movement artefacts or technical reasons.	9 / 38 = 23.7%
Kotilahti et al. (2010)	1.8 d	Nurse observed infant´s behaviour and determined the sleep stage. Movement artefacts were identified using an inclinometer.	-	0 / 13 = 0%
Bortfeld et al. (2009)	6 m – 9 m, mean not reported	Video recorded, behaviour assessed on base of these recordings.	Movement artefacts, hair obstruction, not able to obtain a useful block of data or failure to obtain data from one or more channels.	7 / 28 = 25%
Saitoet al. (2006)	4.4 d	Visual observing during experiment.	-	0 / 20 = 0%
Saito et al. (2006)	4.7 d	Visual observing during eperiment.	-	0 / 20 = 0%

	Mean exclusion percentage: 22.9%

Table 4. Exclusions in auditory studies. Abbreviations: m: month, d: day

Tactile

The number of tactile NIRS studies is still low and there are a lot of different aspects of the modality yet to be investigated. Verriotis et al. (2016) have conducted a study examining the processing of noxious stimuli. The processing of affective touch is a subject that has been investigated more extensively in healthy infants (Kida & Shinohara 2013; Miguel et al. 2019; Jönsson et al. 2018; Pirazzoli et al. 2019). These aspects of the tactile modality are rather different from each other and require different matters to be taken in account considering the experiment protocols.

Studying the processing of unpleasant stimuli such as pain in a delicate population like infants and neonates naturally gives rise to some possible ethical problems - one being choosing an appropriate stimulus. Verriotis et al. (2016) used a clinically required routine blood sample taken with a heel lance as a noxious stimulus in their experiment protocol. The study was conducted to neonates, and the exclusion rates were relatively low 9/36. This kind of ethical considerations don´t play a similar role when a more pleasant stimulus is used.

Affective touch in infants has been studied in variety of age points, mean ages of subjects ranging from 56 days to 228.77 days. All of the studies, except one (Kida & Shinohara 2013), reported the exclusion rates that ranged from 24.2% to 44.8%. Since the processing of affective touch is often investigated so that the infant is awake, fussiness and restlessness of the subject and large movement artefacts are the most usual problems encountered during the imaging session. To minimize infant anxiety during the experiment and to avoid the above-mentioned problems, the infants are often let to familiarize themselves with the experiment room before the actual experiment. The subjects are also usually held in the lap of their caretaker during the experiment, to make it appear less daunting. The small number of studies limits the extent to which studies can be compared to other reviewed

studies, that said, the exclusion rates of the tactile studies seem to have similar order of magnitude compared to the visual studies.

Study	Subject age (mean)	Method of monitoring behaviour	Exclusion criteria	Exclusion rate
Miguel et al. (2019)	228.77 d	Video monitoring was done and subjects that completed at least three successful trials were included.	Fussiness or not having three successful trials.	14 / 49 = 28.6%
Pirazzoli et al. (2019)	160.19 d	Video monitoring and motion detection algorithms used for identifying motion artefacts.	Movement artefacts and disrupted individual trials according to video monitoring. Infant excluded if under 4 valid trials.	8 / 22 = 24.2%
Jönsson et al. (2018)	56 d	Video monitoring and visual monitoring during the experiment.	Restlessness during the imaging session.	13 / 29 = 44.8%
Verriotis et al. (2016)	2 d	Visual monitoring during the experiment.	Large movement artefacts, technical reasons.	9 / 36 = 25%
Kida et al. (2013)	6.6 m	Visual monitoring during the experiment.	NA	NA
				Mean exclusion percentage: 30.7%

Table 5. Exclusions in tactile studies. Abbreviations: m: month, d: day

To conclude, the sensory modality studied and the age of the subjects both have a large impact on the study protocol and how successful it is. Studying older infants enables more complex protocols, which is a factor potentially leading to higher exclusion rates. Nevertheless, for building a comprehensive understanding of the development of sensory processing, it is important that future studies are conducted in a range of age groups and with a greater diversity of different stimuli.

CONCLUSIONS

NIRS imaging has been successfully used to study many different aspects of infant brain function including the sensory processing of visual, auditory and tactile modalities. Most of the results are not

replicated in other studies, which is hopefully remedied in the coming years. NIRS studies often report exclusion rates which can, to some extent, be used to measure how successful the experiment was. Yet, this information must always be weighed against the complexity of the experimental protocol and robustness of the results. All the studies in our review, however, didn´t report the exclusion rate. We suggest that the exclusion criteria should be included in the methods and rejection rate in the results section in future NIRS studies.

The probe features and placement vary considerably. Features of the probe depend on the manufacturer and the available set-up, but we observed that a 2-cm inter-optode interval was commonly used. When describing the probe position, different systems can be used but the most common was the international 10-20 EEG system. The EEG standard is an easily applicable, but relatively imprecise way of describing the position of the probe. Since much more dense grids of optode pairs are still needed to avoid random position errors and account for inter-subject variability in the position of functional areas, considering the resolution the current NIRS imaging devices, using EEG standard system, despite its inaccuracy, is acceptable. However, more precise ways of navigating the probe, such as coupling photogrammetry with infant MRI templates can also be used.

Given the variable dropout rates, it is imperative that researchers monitor participants´ behaviour during the experiment. Video monitoring is an advisable method for this. In addition, we suggest that if the infant is reported to be asleep during the study, a more precise method than visual observing of infant's state is used, e.g., EEG monitoring (as in Verriotis et al. 2016). Otherwise, it may be more appropriate to report that the infants were in an observed sleep or "in a quiet state with their eyes closed" (as Liao et al. 2010). There is evidence that behavioural states such as sleep affect the cortical hemodynamic responses to processing of sensory input (Taga et al. 2018).

Finally, an interesting aspect of future studies will be the use of portable devices to measure hemodynamic brain responses in more natural environments than the experiment room (Paola Pinti

et al. 2018). In addition to portability and easy accessibility, future technological development of instrumentation and analysis methods can enhance the quality and precision of the data captured and also open new possibilities considering multimodal imaging and merging NIRS data with data from other neuroimaging methods.

Acknowledgements & funding

We acknowledge the computational resources provided by the Aalto Science-IT project.

REFERENCES

Bendall, R.C.A., Eachus, P. & Thompson, C., 2016. A Brief Review of Research Using Near-Infrared Spectroscopy to Measure Activation of the Prefrontal Cortex during Emotional Processing: The Importance of Experimental Design. *Frontiers in Human Neuroscience*, 10, pp.1–7.

Bennett, R.H. et al., 2014. fNIRS detects temporal lobe response to affective touch. *Social Cognitive and Affective Neuroscience*, 9, pp.470–476.

Bortfeld, H., Fava, E. & Boas, D.A., 2010. Identifying Cortical Lateralization of Speech Processing in Infants Using Near-Infrared Spectroscopy. *Developmental Neuropsychology*, 34, pp.52–65.

Bortfeld, H. & Wruck, E., 2012. Assessing Infants' Cortical Response to Speech Using Near-Infrared Spectroscopy. *NeuroImage*, 100, pp.130–134.

Brigadoi, S. & Cooper, R. J., 2015. How short is short? Optimum source–detector distance for short-separation channels in functional near-infrared spectroscopy. Neurophotonics, 2, p.1–9.

Ehlis, A.-C. et al., 2014. Application of functional near-infrared spectroscopy in psychiatry. *NeuroImage*, 85, pp.478–488.

Fang, Q. & Boas, D., 2009. Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units. Optics Express, 17, pp. 20178–20190.

Fava, E. et al., 2015. Hemodynamic Responses to Speech and Music in Preverbal Infants. *Child Neuropsychology*, 20, pp.430–448.

Ferradal, S.L. et al., 2014. Atlas-based head modeling and spatial normalization for high-density diffuse optical tomography: In vivo validation against fMRI. *NeuroImage*, 85, pp.117–126.

Fukui, Y., Ajichi, Y. & Okada, E., 2003. Monte Carlo prediction of near-infrared light propagation in realistic adult and neonatal head models. *Applied Optics*, 42, pp.2881–2887.

Hashempour, N. et al., 2019. A novel approach for manual segmentation of hippocampus and

amygdala in newborns' MRI. *Frontiers Neuroscience*, 13, pp.1–15.

Heiskala, J. et al., 2009. Probabilistic atlas can improve reconstruction from optical imaging of the neonatal brain. *Optics Express*, 17, pp.14977–14992.

Honda, Y. et al., 2010. How do infants perceive scrambled face?: A near-infrared spectroscopic study. *Brain Research*, 1308, pp.137–146.

Ichikawa, H. et al., 2018. A longitudinal study of infant view-invariant face processing during the fi rst 3 – 8 months of life. *NeuroImage*, 186, pp.817–824.

Ichikawa, H. et al., 2014. Hemodynamic response of children with attention-de fi cit and hyperactive disorder (ADHD) to emotional facial expressions. *Neuropsychologia*, 63, pp.51–58.

Issard, C. & Gervain, J., 2016. Adult-like processing of time-compressed speech by newborns: A NIRS study. *Developmental Cognitive Neuroscience*, 25, pp.176–184.

Issard, C. & Gervain, J., 2018. Variability of the hemodynamic response in infants: Influence of experimental design and stimulus complexity. *Developmental Cognitive Neuroscience*, 33, pp.182–193.

Jönsson, E.H. et al., 2018. Affective and non-affective touch evoke differential brain responses in 2-month-old infants. *NeuroImage*, 169, pp.162–171.

Kida, T. & Shinohara, K., 2013. Gentle touch activates the prefrontal cortex in infancy: An NIRS study. *Neuroscience Letters*, 541, pp.63–66.

Kobayashi, M. et al., 2011. Do infants represent the face in a viewpoint-invariant manner? Neural adaptation study as measured by near-infrared spectroscopy. *Frontiers in Human Neuroscience*, 5, pp.1-12.

Kobayashi, M. et al., 2014. The processing of faces across non-rigid facial transformation develops at 7 month of age : a fNIRS-adaptation study. *BMC Neuroscience*, 15, pp.1–8.

Kotilahti, K. et al., 2010. Hemodynamic responses to speech and music in newborn infants. *Human*

559 *Brain Mapping*, 31, pp.595–603.
560 Liao, S.M. et al., 2010. Neonatal hemodynamic response to visual cortex activity : high-density
561 near-infrared spectroscopy study. *Journal of Biomedical Optics*, 15, pp.1–9.
562 Lloyd-Fox, S., Blasi, A. & Elwell, C.E., 2010. Illuminating the developing brain: The past, present
563 and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral Reviews*,
564 34, pp.269–284.
565 Makni, S. et al., 2009. Bayesian deconvolution of fMRI data using bilinear dynamical systems.
566 *NeuroImage*, 46, pp.1381.1396.
567 Maria, A. et al., 2018. Emotional Processing in the First 2 Years of Life : A Review of Near-
568 Infrared Spectroscopy Studies. *Journal of Neuroimaging*, 28, pp.441–454.
569 Meek, J., 2002. Basic principles of optical imaging and application to the study of infant
570 development. *Developmental Science*, 5, pp.371–380.
571 Miguel, H.O. et al., 2019. Brain mechanisms for processing discriminative and affective touch in 7-
572 month-old infants. *Developmental Cognitive Neuroscience*, 35, pp.20–27.
573 Minagawa-kawai, Y. et al., 2011. Assessing signal-driven mechanisms in neonates: brain responses
574 to temporally and spectrally different sounds. *Frontiers in Psychology*, 2, pp.1–12.
575 Minagawa-kawai, Y. et al., 2007. Optical Imaging of Infants' Neurocognitive Development: Recent
576 Advances and Perspectives. *Developmental Neurobiology*, 68, pp.712–728.
577 Nakato, E. et al., 2011. Distinct differences in the pattern of hemodynamic response to happy and
578 angry facial expressions in infants - A near-infrared spectroscopic study. *NeuroImage*, 54,
579 pp.1600–1606.
580 Okamoto, M. et al., 2004. Three-dimensional probabilistic anatomical cranio-cerebral correlation
581 via the international 10-20 system oriented for transcranial functional brain mapping.
582 *NeuroImage*, 21, pp.99–111.
583 Pinti, P. et al., 2018. A Review on the Use of Wearable Functional Near-Infrared Spectroscopy in

584 Naturalistic Environments. *Japanese Psychological Research*, 60, pp.347–373.
585 Pinti, P. et al., 2018. The present and future use of functional near-infrared spectroscopy (fNIRS)
586 for cognitive neuroscience. *Annals of the New York Academy of Sciences*, (special issue), pp.1–
587 25.
588 Pirazzoli, L. et al., 2019. Hand or spoon? Exploring the neural basis of affective touch in 5-month-
589 old infants. *Developmental Cognitive Neuroscience*, 35, pp.28–35.
590 Ranger, M. et al., 2011. Cerebral near-infrared spectroscopy as a measure of nociceptive evoked
591 activity in critically ill infants. *Pain Research and Management*, 16, pp.331–336.
592 Saito, Y. et al., 2007. Frontal cerebral blood flow change associated with infant-directed speech.
593 *Archives of disease in childhood. Fetal and neonatal edition*, 92, pp.F113–F116.
594 Saito, Y. et al., 2007. The function of the frontal lobe in neonates for response to a prosodic voice.
595 *Early Human Development*, 83, pp.225–230.
596 Scholkmann, F. et al., 2014. A review on continuous wave functional near-infrared spectroscopy
597 and imaging instrumentation and methodology. *Neuroimage*, 85, pp.6–27.
598 Schroeter, M. et al., 2002. Near-infrared spectroscopy can detect brain activity during a color-word
599 matching Stroop task in an event-related design. *Human Brain Mapping*, 17, pp.61–71.
600 Shekhar, S. et al., 2019. Hemodynamic responses to emotional speech in two-month-old infants
601 imaged using diffuse optical tomography. *Scientific Reports*, 9, pp.1–15.
602 Signe Bray, X.C., 2012. A quantitative comparison of NIRS and fMRI across multiple cognitive
603 tasks. *NeuroImage*, 54, pp.2808–2821.
604 Soltanlou, M. et al., 2018. Applications of Functional Near-Infrared Spectroscopy (fNIRS) in
605 Studying Cognitive Development: The Case of Mathematics and Language. *Frontiers in*
606 *Psychology*, 9, pp.1-15.
607 Sood, B.G., McLaughlin, K. & Cortez, J., 2015. Near-infrared spectroscopy: Applications in
608 neonates. *Seminars in Fetal and Neonatal Medicine*, 20, pp.164–172.

609 Taga, G., Watanabe, H. & Homae, F., 2018. Developmental changes in cortical sensory processing
610 during wakefulness and sleep. *NeuroImage*, 178, pp.519–530.
611 Taga, G., Watanabe, H. & Homae, F., 2011. Spatiotemporal properties of cortical haemodynamic
612 response to auditory stimuli in sleeping infants revealed by multi-channel near-infrared
613 spectroscopy. *Philosophical Transactions of the Royal Society*, 369, pp.4495–4511.
614 Telkemeyer, S. et al., 2011. Acoustic processing of temporally modulated sounds in infants :
615 evidence from a combined near-infrared spectroscopy and EEG study. *Frontiers in*
616 *Psychology*, 2, pp.1–14.
617 Urakawa, S. et al., 2015. Selective Medial Prefrontal Cortex Responses During Live Mutual Gaze
618 Interactions in Human Infants: An fNIRS Study. *Brain Topography*, 28, pp.691–701.
619 Verriotis, M. et al., 2016. Mapping Cortical Responses to Somatosensory Stimuli in Human Infants
620 with Simultaneous Near-Infrared Spectroscopy and Event-Related Potential Recording.
621 *eNeuro*, 3, pp.1–15.
622 White, B.R. & Culver, J.P., 2010. Quantitative evaluation of high-density diffuse optical
623 tomography: in vivo resolution and mapping performance. *Journal of Biomedical Optics*,
624 15(2), pp. 026006-1–026006-13.
625 Wilcox, T. et al., 2013. Functional Activation of the Infant Cortex during Object Processing.
626 *Neuroimage*, 62, pp.1833–1840.
627 Wilcox, T. et al., 2007. Using near-infrared spectroscopy to assess neural activation during object
628 processing in infants. *Journal of Biomedical Optics*, 10, pp.1–16.
629 Wilcox, T., Bortfeld, H. & Woods, R., 2008. Hemodynamic Response to Featural Changes in the
630 Occipital and Inferior Temporal Cortex in Infants: A Preliminary Methodological Exploration.
631 *Developmental Science*, 11, pp.361–370.
632 Yang, J. et al., 2016. Cortical response to categorical color perception in infants investigated by
633 near-infrared spectroscopy; SUPPORTING INFORMATION. *PNAS*, 113(9).

634 Yang, J. et al., 2016. Cortical response to categorical color perception in infants investigated by
635 near-infrared spectroscopy. *PNAS*, 113, pp.2370–2375.
636 Yao, R., Intes, X. & Fang, Q., 2018. Direct approach to compute Jacobians for diffuse optical
637 tomography using perturbation Monte Carlo-based photon "replay". Biomedical Optics Express
638 9, pp.4588–4603.

639

640